

П.Ф. ЩАПОВ, Т.Г. ОСИНА, ГАЙДАШ А.М. (г. Харьков)

ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ РЕШАЮЩИХ ПРАВИЛ ПРИ ЭКСПРЕСС-КОНТРОЛЕ НЕОПРЕДЕЛЕННЫХ ПАРАМЕТРОВ СЫПУЧИХ МАТЕРИАЛОВ

У статті розглянуто результати використання метода параметричної дискримінації рівній клейковини насіння пшениці в мовах виключення вектору вхідних вимірювальних сигналів системи контролю.

In the article the results of the use of method of self-reactance discrimination of levels of gluten of grain of wheat on condition of exception of vector of entrance measuring signals of the checking system are considered.

Постановка проблемы. Основные трудности, возникающие при оценке параметров, характеризующих состав, например, биологически сложных сыпучих материалов, связаны с отсутствием не только стандартных образцов, воспроизводящих заданные значения, но и с невозможностью прямых измерений для этих параметров. Если рассматривать такие комплексные параметры, как качество и количество клейковины в зерне пшеницы, то эти трудности усугубляются влиянием человеческого фактора [1], снижающим объективность результатов количественного и качественного анализа. Этот же фактор присутствует и при измерении уровня масличности семян подсолнечника.

Анализ литературы. Несмотря на очевидные преимущества и распространенность методов и технических средств экспресс-контроля следует отметить принципиальные сложности и недостатки такого контроля, особенно если речь идет об измерительном контроле качественных параметров сыпучих сырьевых материалов.

Подавляющее число методов экспресс-контроля сыпучих материалов основано на косвенных измерениях параметров качества, что не позволяет устранять мешающее влияние неконтролируемых факторов (насыпная плотность, гигроскопичность, когезионные и адгезионные свойства и т.д.), уровни которых не только плохо поддаются нормировке, но и зависят от меняющихся условий измерительного эксперимента [2].

Многopараметрический экспресс-контроль отличается пониженной достоверностью, поскольку информационные сигналы используемых измерительных преобразователей взаимно коррелированы, что снижает количество суммарной ожидаемой измерительной информации [3].

Метрологическое обеспечение измерений при контроле технологических параметров сыпучих материалов со сложной биохимической структурой оставляет желать лучшего, поскольку однозначное воспроизведение уровней таких параметров при обучении (градуировке) системы контроля

затруднено из-за неопределенности физико-механических свойств сыпучих материалов, вызывающих статистическую неоднородность результатов измерений [4].

Цель статьи – статистическое обоснование применения методов многомерной дискриминации уровней априорно неопределенных параметров органических сыпучих материалов.

Общая модель принятия решений. Пусть объект контроля характеризуется p -мерным вектором $X = (X_1, \dots, X_p)$, составляющие которого являются информационными признаками (единичными показателями контроля). В этом случае вектор X можно рассматривать как комплексный показатель. Предполагается, что распределение признаков X_1, \dots, X_p для каждого класса π_1, \dots, π_q состояний объекта контроля можно описать условной функцией плотности распределения вероятностей $f(X_i / \pi_r), i = 1, p; r = 1, q$.

Если цены неправильных решений как первого, так и второго рода неизвестны, а априорные вероятности классов π_1, \dots, π_q одинаковы, то в соответствии с байесовым принципом минимизации среднего риска [5], решение о выборе класса принимается по максимуму функции правдоподобия

$$F(x_1, \dots, x_p / \pi_r) = \prod f(x_i / \pi_r),$$

где x_1, \dots, x_p – результаты измерения значений единичных показателей по всем соответствующим векторам \bar{X} .

Для случая двух классов (π_1 и π_2) используют дискриминантную (решающую) функцию, представляющую логарифм отношения правдоподобия

$$g(x) = \sum [\ln F(x_1, \dots, x_p / \pi_1) - \ln F(x_1, \dots, x_p / \pi_2)]. \quad (1)$$

Принятие одного из двух возможных решений сводится к тестированию выборки (x_1, \dots, x_p) , где критериальной статистикой является $g(x)$, а выбор решений $\gamma_1, (\pi \in \pi_1)$, или $\gamma_2, (\pi \in \pi_2)$, принимают в соответствии с правилом выбора:

$$\begin{cases} \gamma_1 : g(x) \geq 0, \\ \gamma_2 : g(x) < 0. \end{cases}$$

Если условные плотности $f(x_i / \pi_r)$ результатов измерения независимых составляющих X_1, \dots, X_p имеют нормальный закон распределения со средними значениями $\mu_i^{(r)}$, дисперсиями $\sigma_i^2, i = \overline{1, p}; r = \overline{1, 2}$, то дискриминантная функция (1) имеет вид [7]

$$g_2(x) = [X - 0,5(\bar{x}^{(1)} - \bar{x}^{(2)})] D^{-1} (\bar{x}^{(1)} - \bar{x}^{(2)}), \quad (2)$$

где $\bar{x}^{(1)}, \bar{x}^{(2)}$ – оценки средних, D – оценка дисперсии.

Таблица 1 – Результаты регрессионного анализа многомерных измерений

№	Уравнение с коэффициентами	P	R ²	Стандартная ошибка, %	F _{p,1,p-9} расчет	\bar{R}_p^2	Нормирован- ный R ²
1	$y = 33,09717084 - 0,822683501 \cdot X_1$	2	0,158115081	3,473418574	22,16167451	0,143845845	0,150980463
2	$y = 30,71424 - 0,22736 \cdot X_2$	2	0,06671	3,657118	8,434475	0,050892	0,058801
3	$y = 30,05615 - 0,40261 \cdot X_3$	2	0,061161	3,667975	7,687098	0,045248	0,053204
4	$y = 39,97278 - 6,3313 \cdot X_4$	2	0,040138	3,708816	4,934286	0,023869	0,032003
5	$y = 38,6644401 - 0,753722051 \cdot X_1 - 0,291838671 \cdot X_3$	3	0,19951	3,401392	14,58027	0,178985	0,185827
6	$y = 37,56509 - 0,72882 \cdot X_1 - 0,1369 \cdot X_3$	3	0,180241	3,442087	12,86245	0,159222	0,166228
7	$y = 52,41218 - 0,82597 \cdot X_1 - 6,42992 \cdot X_4$	3	0,189315196	3,422983821	13,66121	0,168528	0,175457336
8	$y = 31,30758 - 0,14963 \cdot X_2 - 0,17418 \cdot X_3$	3	0,070341	3,665564	4,426269	0,046503	0,054449
9	$y = 31,43604 - 0,22009 \cdot X_2 - 0,34457 \cdot X_4$	3	0,066761	3,672614	4,184899	0,042832	0,050808
10	$y = 22,999 - 0,57049 \cdot X_3 + 3,609656 \cdot X_4$	3	0,063667	3,678697	3,977768	0,039658	0,047661
11	$y = 38,71142112 - 0,750209838 \cdot X_1 - 0,276691443 \cdot X_3 - 0,010413434 \cdot X_2$	4	0,189356872	3,437618034	9,032086757	0,161403661	0,1688391964
12	$y = 54,00207 - 0,84853 \cdot X_1 + 0,032234 \cdot X_2 - 7,30941 \cdot X_4$	4	0,199963	3,415057	9,664406	0,172375	0,179272
13	$y = 23,43186 - 0,15505 \cdot X_2 - 0,35414 \cdot X_3 + 4,051524 \cdot X_4$	4	0,073486	3,675097	3,066829	0,041537	0,049524
14	$y = 55,78698 - 0,85063 \cdot X_1 + 0,099706 \cdot X_3 - 8,19187 \cdot X_4$	4	0,200105	3,414753	9,67301	0,172522	0,179418
15	$y = 56,09012 - 0,86003 \cdot X_1 + 0,021175 \cdot X_2 + 0,078302 \cdot X_3 - 8,3826 \cdot X_4$	5	0,200271	3,42921	7,199698	0,165501	0,172455

Сравнение моделей непрямого измерения (регрессионной и линейной дискриминации). Для оценки эффективности применения измерительных процедур на базе линейной дискриминации контролируемых уровней клейковины были построены регрессионные модели функции преобразования и выбрана оптимальная (по максимуму скорректированного коэффициента детерминации \bar{R}_p^2 [7]) модель множественной регрессии.

Одновременно с вычислением \bar{R}_p^2 и F -статистики значимости регрессии ($F_{p-1, n-p}$, где p – число параметров регрессии, а n – объем выборки), вычислялась стандартная ошибка σ_e , как среднеквадратическое отклонение случайного остатка регрессионной модели. Вычислялся и нормированный коэффициент множественной корреляции R^2 .

В таблице 1 приведены результаты расчета коэффициентов уравнений множественной регрессии y (клейковины), предикторы которых сформированы на основе четырех показателей контроля:

- X_1 – влажность;
- X_2 – масса 1000 семян;
- X_3 – теплофизический показатель;
- X_4 – линейный размер (в) зерновки.

Как видно из таблицы, уравнение регрессии № 5 ($\bar{R}_p^2 = 0,178985$, $\sigma_e = 3,401392\%$), имеет не только максимум по \bar{R}_p^2 , но и минимальную стандартную ошибку (в процентах содержания клейковины). Учитывая нормальный закон распределения результатов измерения y_i уровней клейковины ($i = 1, 120$), можно вычислить величину достоверности при контроле двух, трех и четырех уровней клейковины ($\mu = 2, 3, 4$).

В таблице 2 представлены результаты расчетов достоверности контроля с использованием описанных четырех показателей $X_1 \dots X_4$:

- а) функциональной (регрессионной) модели измерительного преобразования;
- б) классификационной модели измерительного преобразования (модель дискриминации).

Таблица 2

μ	Достоверность контроля	
	регрессионная модель	дискриминационная модель
2	0,702	0,722
3	0,680	0,754
4	0,669	0,725

Из таблицы 2 видно, что с увеличением числа контролируемых уровней достоверность контроля функциональной модели падает. Она эффективна при $\mu=2$, но при $\mu=3$ и, тем более $\mu=4$, более эффективна классификационная модель. Результаты статистического анализа показали, что СКО измерения (для $\mu=3$ и $\alpha=\beta$) равна, для классификационной модели, величине $\sigma_{\Delta}=1,99\%$, что меньше, чем $\sigma_e = 3,40\%$.

Практические результаты. Описанная система контроля была использована в экспресс-контроле содержания клейковины в зерне на предприятии "Инженерная группа "ТФК", что позволило решить задачу оптимизации процессов дозирования зерна при измерении содержания клейковины по трем контролируемым уровням.

Выводы. Проведен сравнительный анализ регрессионной и классификационной моделей измерительного преобразования при физическом моделировании многопараметрической системы контроля уровня количественного содержания клейковины в зерне пшеницы. Подтверждено преимущество модели линейной дискриминации по сравнению с функциональными (регрессионными) моделями измерительных преобразований в задаче повышения достоверности контроля количества клейковины зерна пшеницы.

Список литературы: 1. Данильчук П.В., Торжиская Л.Р. Оценка качества зерна в хозяйствах и на хлебоприемных предприятиях: Справочник. – К.: Урожай, 1990. 2. Шубин И.Н., Свиридов М.М., Таров В.П. Технологические машины и оборудование. Сыпучие материалы и их свойства: Учебн. пособие. – Тамбов: Изд-во ТГТУ, 2005. – 76 с. 3. Щапов П.Ф., Осина Т.Г. Оценка информационной значимости показателей измерительного контроля метрологических неопределенных параметров зерна пшеницы. Сб. науч. труд. Том VII. МНК «Метрология и измерительная техника». – Харьков: ХНУРЭ, 2005. – С. 146-149. 4. Щапов П.Ф., Качанов М.П., Гайдаш А.М. Анализ факторного влияния влажности и плотности сыпучих материалов на информационные параметры резонансного преобразователя // Вісник НТУ «ХПІ». Збірник наукових праць. Тематичний випуск: Автоматика та приладобудування. – 2005. – №7. – С. 167-172. 5. Осина Т.Г. Использование параметрических дискриминантных функций в задачах измерительного контроля качества // Вестник Национального Технического Университета «ХПИ». – Харьков: НТУ «ХПИ». – 2006. – вып. 9. – С. 103-107. 6. Себер Дж. Линейный регрессионный анализ. – М.: Мир, 1980. – 456 с.

Поступила в редакцию 18.06.08